

## **CNN-SVO: improving the mapping in semi-direct visual odometry using single-image depth prediction**

### **ABSTRACT**

Reliable feature correspondence between frames is a critical step in visual odometry (VO) and visual simultaneous localization and mapping (V-SLAM) algorithms. In comparison with existing VO and V-SLAM algorithms, semi-direct visual odometry (SVO) has two main advantages that lead to state-of-the-art frame rate camera motion estimation: direct pixel correspondence and efficient implementation of probabilistic mapping method. This paper improves the SVO mapping by initializing the mean and the variance of the depth at a feature location according to the depth prediction from a single-image depth prediction network. By significantly reducing the depth uncertainty of the initialized map point (i.e., small variance centred about the depth prediction), the benefits are twofold: reliable feature correspondence between views and fast convergence to the true depth in order to create new map points. We evaluate our method with two outdoor datasets: KITTI dataset and Oxford Robotcar dataset. The experimental results indicate that improved SVO mapping results in increased robustness and camera tracking accuracy. The implementation of this work is available at <https://github.com/yan99033/CNN-SVO>.